

The Specific Relation Between Perception and Production Errors for Place of Articulation in Developmental Apraxia of Speech

Paul Groenen

University Hospital Nijmegen
Department of Otorhinolaryngology
Child Neurology Center
Institute of Medical Psychology
Nijmegen, The Netherlands

Ben Maassen

University Hospital Nijmegen
Child Neurology Center
Institute of Medical Psychology
Nijmegen, The Netherlands

Thom Crul

University Hospital Nijmegen
Department of Otorhinolaryngology
Institute of Medical Psychology
Nijmegen, The Netherlands

Geert Thoonen

University Hospital Nijmegen
Child Neurology Center
Institute of Medical Psychology
Nijmegen, The Netherlands

Developmental apraxia of speech is a disorder of phonological and articulatory output processes. However, it has been suggested that perceptual deficits may contribute to the disorder. Identification and discrimination tasks offer a fine-grained assessment of central auditory and phonetic functions. Seventeen children with developmental apraxia (mean age 8:9, years:months) and 16 control children (mean age 8:0) were administered tests of identification and discrimination of resynthesized and synthesized monosyllabic words differing in place-of-articulation of the initial voiced stop consonants. The resynthetic and synthetic words differed in the intensity of the third formant, a variable potentially enlarging their clinical value. The results of the identification task showed equal slopes for both subject groups, which indicates no phonetic processing deficit in developmental apraxia of speech. The hypothesized effect of the manipulation of the intensity of the third formant of the stimuli was not substantiated. However, the children with apraxia demonstrated poorer discrimination than the control children, which suggests affected auditory processing. Furthermore, analyses of discrimination performance and articulation data per apraxic subject demonstrated a specific relation between the degree to which auditory processing is affected and the frequency of place-of-articulation substitutions in production. This indicates the interdependence of perception and production. The results also suggest that the use of perceptual tasks has significant clinical value.

KEY WORDS: developmental apraxia of speech (DAS), speech perception, speech production, place of articulation, children

The notion that developmental apraxia of speech (DAS) is a pathological syndrome has been abandoned (Guyette & Diedrich, 1981). Instead, DAS is considered to be a developmental disorder in the ability to perform purposeful speech movements (Hall, Jordan & Robin, 1993). The symptoms include frequent and inconsistent errors that are contextually constrained in the articulation of consonants and vowels (Crary, 1984; Crary, Landess, & Towne, 1984; Marquardt, Dunn, & Davis, 1985; MiCoCh & Square, 1984; Stackhouse, 1992). However, the explanation of DAS as an output speech disorder does not preclude the possibility that most (but not all) children demonstrate additional language or language-related problems (Bridgeman & Snowling, 1988; Ekelman & Aram 1983; Hall et al., 1993; Marion, Sussman, & Marquardt, 1993; Snowling & Stackhouse, 1983).

In the present study, the relation between the perception and production of speech in the etiology and maintenance of DAS was examined. Several studies have demonstrated a relation between speech perception and articulatory deficits in

diverse groups of subjects. Hoffman, Daniloff, Bengoa, and Schuckers (1985) and Ohde and Sharf (1988) found that children with articulation problems had problems distinguishing /r/ from /w/ consistently. In addition, Monnin and Huntington (1974) found a specific relation between identification and production of /r/-/w/ contrasts in subjects with speech defects. A specific relation between perception and production in children with articulation problems was also found by Raaymakers and Crul (1988). They studied the final /s-ts/ contrast. Rvachew and Jamieson (1989) investigated the perception of fricatives in children with a functional articulation disorder. Their results suggested that for a subgroup of children with functional articulation disorders, production errors may reflect speech perception errors.

Few studies have been conducted on perception and production skills in a specific group of children with developmental apraxia of speech. Hoit-Dalgaard, Murry, and Kopp (1983) studied voice onset time production and perception in subjects with apraxia. They found that children with apraxia had production as well as perceptual errors for the voicing feature. They did not, however, find a significant relationship between the two measures. Marion et al. (1993) found subjects with apraxia to have rhyming abilities inferior to those shown by control children. Hence, little is known about the relation between the perception and production of specific phonetic features in children with DAS.

According to Locke (1980a, 1980b), tests of perception with phonemes related to the specific articulation errors of the child may be particularly useful for the assessment of the relation between speech perception and production. Thoonen, Maassen, Gabreëls, and Schreuder (1994) identified a group of children with apraxia who specifically made errors on the initial consonantal place-of-articulation. For this reason, our primary interest in the present study is in the specific relation between the perception and production of place-of-articulation for consonants in children with developmental apraxia of speech.

Data on the auditory perceptual skills of children with DAS are scarce (Hall et al., 1993). Guyette and Diedrich (1981) along with Edwards (1984) also explicitly mention the inconclusiveness of the available auditory findings. This inconclusiveness may reflect the use of inappropriate or insensitive test procedures (Locke, 1980a). It may also reflect the lack of differentiation of speech perception into distinct processes.

In the present study, a sufficiently sensitive test for the perception of speech was developed using speech sound continua. A speech continuum consists of a series of speech tokens that vary acoustically for a single phonological contrast. The redundancy of the speech signal used in this study was reduced by focusing on acoustic cues for a single dimension (i.e., the second and third formant transitions for cuing place of articulation).¹

Speech perception can be characterized by a series of processes, including a preliminary auditory analysis, further

auditory and phonetic feature analysis, and the combination of phonetic features into a phonemic representation (Cutting & Pisoni, 1978; Pisoni & Sawusch, 1975). At any stage in this process, information can be placed in short-term memory. Auditory processing includes a preliminary analysis and is related to auditory short-term memory, whereas phonetic processing includes phonemic labeling strategies and is related to phonetic memory (Baddeley, 1992). In order to distinguish the different processes in the auditory perception of speech, identification and discrimination tasks were used.

An identification task requires a phonemic judgment, and thus decisions are based primarily on the phonetic properties and features represented in phonetic short-term memory. The decisions involved in a discrimination task may be based on information from both phonetic and auditory memory. (See Liberman, Harris, Hoffman, & Griffith, 1957; and Studdert-Kennedy, Liberman, Harris, & Cooper, 1970, for the early studies on categorical perception.)

A secondary interest in the present study is related to the differential effects of lowering the intensity of the third formant by using natural versus synthetic speech. The structure of the third formant is viewed to be a minor cue for the Dutch place-of-articulation contrasts. We wanted to compare the sensitivity of listeners to a reduction in the high frequency information of the signals, specifically to the minor cue for the b-d dimension. In our resynthesized natural speech all three formants (F1, F2, and F3) are well preserved, whereas in our synthetic speech the intensity of the third formant is decreased. Huntress, Lee, Creaghead, Wheeler, and Braverman (1990) found that aphasic subjects comprehended natural speech better than synthetic speech, particularly when the subjects first heard the synthetic speech. More than half of the subjects in the study of Huntress et al. demonstrated mild-to-severe apraxia of speech, and the authors did not compare the subjects to a control group. Our synthetic stimuli may be more sensitive in the detection of perception problems than natural stimuli because of a lack of redundancy in the signal and because there was lowered intensity in the third formant frequency. Repp (1984) pointed to naturalness as a stimulus factor in normal speech perception and suggested that impoverished speech stimuli may impoverish perception. Natural speech is more acoustically redundant than synthetic speech. Thus, the question is whether children with apraxia need redundancy in the speech signal as in the current investigation's "resynthesized" natural tokens. Furthermore, it was questioned whether lowering the intensity of the third formant of the synthetic tokens would differentially influence the perceptual behavior of the group with apraxia compared to the control group.

In Experiment 1, we focused on the perceptual abilities of a group with apraxia and a control group in response to a place-of-articulation continuum. Both the children with apraxia and the control children performed identification and discrimination tasks. In Experiment 2, the perceptual functions of the children with apraxia were compared to their specific articulatory behavior with a focus on place-of-articulation.

¹Important acoustic cues carrying place information are the transitions of the second and third formants and the spectrum of the release burst. We decided to manipulate the formant transitions.

EXPERIMENT 1

Method

Subjects

The purpose was to form a homogeneous group of children whose main problem was apraxic in nature. The subjects with apraxia were 17 children (mean age 8:9; range 6:11 to 11:06) attending special schools for children with language and speech disorders in three Dutch cities.

In the preselection, medical and educational records were examined, and the school speech-language pathologists performed speech evaluations. The criteria for DAS were derived from the characteristics mentioned in Hall (1992), Hall et al. (1993), and Stackhouse (1992). The criteria for inclusion were (a) high rates of speech sound errors, (b) inadequate diadochokinetic profile for the production of multisyllabic sequences (e.g., /pataka-pataka-pataka . . . /), (c) posturing and groping of the articulators, (d) periods of highly unintelligible speech, (e) difficulties with or inability to produce complex phonemic sequences, (f) high incidences of context-related sound substitutions (e.g., metathetic errors), and (g) an inconsistent speech performance. An inclusion criterion obtained from the medical and educational records was slow development and remediation of speech skills.

In addition, information derived from the medical and educational records was used to determine exclusion criteria (see also Thoonen et al., 1994; Thoonen, Maassen, Wit, Gabreëls, & Schreuder, in press). This information indicated that each selected child (a) had no structural problems in the speech organs that could be held responsible for their speaking problems, (b) did not have otorhinolaryngologic problems, and (c) did not suffer from severe attention deficits. Each child with DAS functioned within a normal range of intelligence (IQ-range on standardized tests of intelligence was 84 to 108).

After the preselection, each child was tested during a short screening session. This screening included (a) imitation of pitch and temporal changes while sustaining the vowel /a/, (b) 10 minutes of spontaneous speech, and (c) imitation of eight short sentences. The first task (imitation of pitch and temporal changes) was used to determine whether the child could comprehend specific task instructions, and the latter two tasks were used to ascertain that each child had a complete phonemic repertoire.

The final admission to the subject groups required the child to pass another phase of selection. Each child had to be unequivocally diagnosed by certified speech-language pathologists (Child Neurology Centre, University Hospital Nijmegen) as a clear case of DAS. Audiotaped recordings of spontaneous speech and sentence imitations were used for diagnostic classification. The speech-language pathologists had no access to the medical and educational records. The speech criteria used for the classification of DAS were identical to those used in the preselection by the school speech-language pathologists. Besides developmental apraxia of speech, the speech-language pathologists also screened for dysarthria, dysphasia, functional articulation

problem, and speech-language delay. Admission to the experimental DAS group required a perfect (100%) agreement between the two speech-language pathologists in categorizing the speech characteristics as moderate to severe symptoms of DAS.

The control subjects were 16 children (mean age 8:0; range 6:04 to 10:02) attending a regular elementary school and recommended by their teachers. The children did not evidence learning disabilities, a history of hearing problems, speech and language problems, or speech-limiting structural abnormalities. On the basis of school performance and information from the classroom teachers, normal levels of cognitive, motoric, and perceptual functioning could be assumed. The control children were gender-matched to those with DAS and were in the same school grade, so the educational level was the same across groups. The mean age of the control subjects was slightly younger than that of the subjects with apraxia.

The children in both groups also met the following selection criteria: (a) absence of hearing loss on bilateral pure tone audiometric testing with air-conduction thresholds at 250, 500, 1000, 2000, 4000 Hz (ISO, 1985); the maximally allowed hearing loss was 25 dB HL for either ear; (b) no previous exposure to resynthesized or synthesized speech; and (c) Dutch as the native language. In addition, only children who could correctly identify 11 out of a series of 12 words consisting of six random repetitions of two speech tokens representing the perceptually clearest ends of the speech continua (i.e., /bæk/ and /dæk/) were admitted to the study. The probability of obtaining 11 correct responses out of 12 trials based on chance alone was .003. The subjects had to pass this pretest so that we could exclude children who had difficulties accommodating to artificial speech. The pretest was administered before actual testing, and all of the 17 children with apraxia in the experimental group and 16 children in the control group passed the pretest.

Stimuli

Two seven-step /b-d/ continua were generated. The first continuum was based on a natural adult male voice. By manipulation of the linear predictive coding parameters and resynthesis of the result, a "resynthetic" natural continuum was constructed. The starting point was utterance of the single word /bæk/ (the Dutch word for *box*). After A/D conversion with a DASH-16 data-acquisition board (12 bit sampling at 10 kHz; band-pass filtering between 40 and 5000 Hz, low-pass cut-off frequency 5000 Hz with a decline of 60 dB/octave), the Interactive Laboratory System (ILS, V6.1, 1989) was used to manipulate the spectral structure of the initial formant transitions. Only the vowel portion (formant transitions plus steady-state vowel) was analyzed with pitch-synchronous linear predictive coding (covariance method: pre-emphasis factor .98, Hamming window), which yielded 12 reflection coefficients (Markel & Gray, 1976). The locations of the spectral peaks, their bandwidths, and their intensities were estimated by transforming the reflection coefficients to autoregressive coefficients and then performing a fast Fourier transformation (FFT). The result was

TABLE 1. Onset frequencies (in Hz) for the second and third formant transitions from the resynthetic and synthetic stimuli.

Stimulus	F2	F3	
1	1000	2150	/bæk/
2	1083	2317	
3	1167	2483	
4	1250	2650	
5	1333	2817	
6	1417	2983	
7	1500	3150	/dæk/

smoothed spectrally by interactively adjusting the formant frequencies.

The consecutive stimuli of the resynthetic continuum ranged perceptually from /bæk/ to /dæk/ (i.e., the Dutch word for *roof*) and differed from one another in the starting value and slope of the transitions of the second and third formant. The onset frequencies of the F2 and F3 for each stimulus are shown in Table 1. F1 always started at 400 Hz. The transition of the first formant was 20 ms in duration. The transitions of the second and third formants were 52 ms in duration. All transitions were linear. The final 98 ms of the vowel consisted of steady-state formants appropriate for the Dutch vowel /a/ with center frequencies at 750 Hz (F1), 1150 Hz (F2), and 2500 Hz (F3).

The sampled data were resynthesized with a pitch-synchronous synthesis procedure by transforming the manipulated reflection coefficients to inverse filter coefficients. Pitch period excitation used a unit pulse. The resynthesized vowel parts were spliced back into the original utterance /bæk/, which produced seven stimuli ranging from /bæk/ to /dæk/. The total length of each stimulus was 381 ms, consisting of (a) voice-lead 71 ms; (b) burst 10 ms; (c) vowel /a/ 150 ms, divided into 52 ms transition duration and 98 ms steady state; (d) silence interval (occlusion period /k/) 70 ms; and (e) release /k/ 80 ms. Stimulus phonemic quality was checked by having 10 adults label 10 repetitions of each of the seven stimuli presented in random order. The phoneme boundary was located for individual subjects between stimulus 3 and 5, indicating a valid choice of frequencies in the spectral manipulations.

The second continuum was a "synthetic" /b-d/ continuum based on a text-to-speech system that followed the principles of allophonic synthesis. As in the resynthetic stimulus set, the consecutive stimuli of the synthetic continuum differed from one another in the starting value and slope of the transitions of the second and third formant (see Table 1).

The construction of the synthetic speech was intended to reduce the intensity of the third formant when compared to the resynthetic speech. This was accomplished by using the allophone-based text-to-speech system developed at the Institute of Phonetics, University of Nijmegen, The Netherlands. This system combines knowledge of the phonetic structure of phonemes with phonetic rules that adjust a phoneme's target values to its contextual surroundings (Boves, Kerkhoff, & Loman, 1987; Loman, Kerkhoff, & Boves, 1989). The linguistic component converts text to a phonemic representation supplemented by suprasegmental informa-

tion about intonation, syllabic structure, and emphasis. The phonetic component interprets the phonemic representation in terms of acoustic patterns, makes phonetic adjustments, and excites a set of serial resonators. It is possible to change each of the parameters before actual synthesis. In order to make a reliable comparison between the resynthetic and the synthetic continuum, the synthetic stimuli were constructed using the temporal and spectral specifications for the resynthetic stimuli, which are linguistically relevant for the place-of-articulation contrast (formant transitions and vowel characteristics). The remaining parameters were determined using the rules of the system, resulting in minor spectral and temporal differences with regard to the resynthetic stimuli.

The spectral structure of the synthetic and resynthetic speech was indicated by long-term average spectra of the vowel part (formant transition and steady-state vowel). The vowel parts from both the resynthetic and synthetic endpoint tokens /bæk/ and /dæk/ were submitted to a series of 128 points FFTs (frame 12.8 ms, pre-emphasis .98, Hamming window). Long-term average spectra were computed by averaging the intensity levels in each frequency band over time. Figures 1 and 2 present the long-term average spectra of the endpoint tokens for /bæk/ and /dæk/, respectively. The most important information from the high and low frequencies in the long-term average spectra can be extracted by modifying two of the five indices originally proposed by Hammarberg, Fritzell, Gauffin, Sundberg, and Wedin (1980). The two indices concern the slope of the spectral envelope (a) below 400 Hz, and (b) between 1.5 and 5 kHz. In the present study, we were not interested in the slope value but in the overall frequency level indicated by the intercept (a) in the regression function. In Figures 1 and 2, the lines labeled with squares represent the results of linear regression analyses on the selected frequency regions.

In the frequencies below 400 Hz, the synthetic speech has a higher intensity than the resynthetic speech. In the frequencies above 1500 Hz, the resynthetic speech has a higher intensity than the synthetic speech. For both continua, frequencies of the third formant were well within the spectral range after 1500 Hz, whereas frequencies of the second formant were within the spectral range before 1500 Hz. Thus, the resynthetic stimuli showed pronounced third formant amplitude, whereas the synthetic stimuli showed decreased intensity levels of the third formant, confirming the validity of our experimental manipulation.

The synthetic and resynthetic stimulus sets were comparable with regard to the spectral cues for place-of-articulation. As a result of different systems for stimulus generation, there were minor acoustic differences between the resynthetic and the synthetic speech (voice-lead: 71 ms–65 ms; occlusion /k/: 70 ms–75 ms; release /k/: 80 ms–30 ms; duration F1-transition: 20 ms–25 ms; duration F2/F3-transition: 52 ms–50 ms; total length stimulus: 381 ms–330 ms, respectively). Hence, these differences did not influence the perception of the feature place-of-articulation.

In Figures 3 and 4, the oscillograms and spectrograms displaying the endpoint tokens for /bæk/ and /dæk/ when resynthesized and synthesized are presented.

Procedure

The stimuli were recorded and played back using an Ampex 467 DAT-tape on a Grundig Fine Arts Digital Audio Tape recorder (type DAT-9000: 16 bit D/A converter, 2-fold oversampling, sampling frequency 48 kHz). Presentation was via a Beyerdynamic closed headphone (Type DT770). The playback level was set at a listening level judged by the subject to be comfortable (approximately 70 dB HL). Subjects were tested in a quiet room.

Each subject was examined twice within a 2-day period. In order to get accustomed to the artificial speech, the subject first listened to six repetitions of each of the four endpoint stimuli. The subject then had to identify 11 out of a series of 12 endpoint stimuli correctly for both the resynthetic and the synthetic speech (see subject selection criteria). After this, the subject was administered the main experimental tasks: two identification and two discrimination tasks. The two synthesized stimuli tasks (identification and discrimination) and the two resynthesized stimuli tasks (identification and discrimination) were presented together, with the order counterbalanced across subjects over the 2 days. The discrimination and identification tasks for one stimulus set was presented on the same day.

The identification task was based on a two-alternative forced-choice response procedure and consisted of 10 repetitions of each of the seven stimuli presented in a random order in five series of 14 stimuli. The stimuli were separated by an interstimulus interval of 5000 ms. The subjects could identify the stimulus by pointing to one of two pictures: a picture of a box, representing the stimulus /bak/, or a picture of a roof, representing the stimulus /dax/. Half of the subjects began with the resynthetic stimuli, whereas the other half began with the synthetic stimuli. The children were encouraged to respond but

never received differential feedback for particular responses.

The discrimination tasks consisted of same-different (AX) judgments for resynthetic and synthetic stimulus pairs. In order to obtain a bias-free measure of discriminability, the discrimination tasks were set up in such a way that signal detection analysis could be applied (Coombs, Dawes, & Tversky, 1970). For this, each task contained physically different as well as identical pairs. In both the resynthetic and synthetic discrimination tasks, the subjects were presented three series of 15 stimulus pairs. Each series consisted of one identical pair for each of the seven stimuli, one physically different pair for each of the five possible 2-step comparisons (1-3, 2-4, 3-5, 4-6, 5-7), and one physically different pair for three 3-step comparisons (2-5, 3-6, 4-7). The latter were treated as dummies explicitly intended to elicit "different" responses. In each series, the pairs were randomly ordered with an intrapair interstimulus interval of 600 ms and an interpair interval of 5000 ms. The subjects were required to point to one of two pictures after hearing a pair of stimuli: a picture of a triangle and a circle, representing the concept "different," or a picture of two circles, representing the concept "same." Half of the subjects began with the resynthetic stimuli, whereas the other half began with the synthetic stimuli.

Results

Identification

In Figure 5, the mean identification curves for the children with apraxia and the control children for the resynthetic and the synthetic stimuli are presented.

Each individual identification curve was submitted to probit transformations (Finney, 1971). The probit method

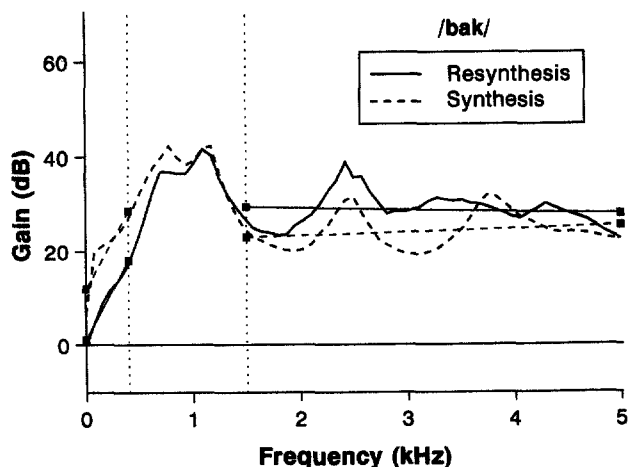


FIGURE 1. Long-term average spectra of the resynthetic and synthetic endpoint tokens for /bak/. The unlabeled lines represent overall intensity levels from 0 to 5000 Hz. The lines labeled with squares represent the results of linear regression analyses on selected frequency regions.

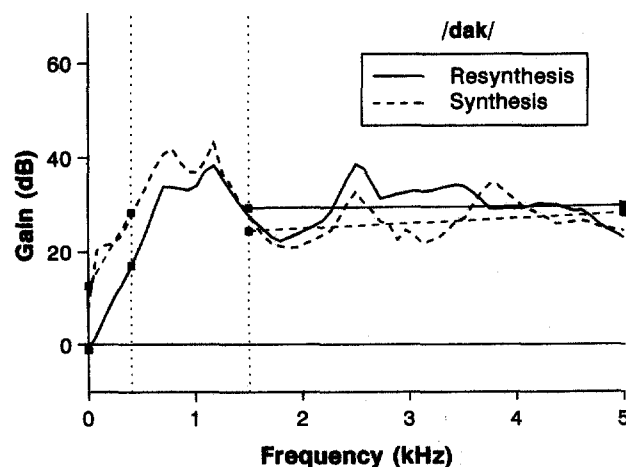


FIGURE 2. Long-term average spectra of the resynthetic and synthetic endpoint tokens for /dak/. The unlabeled lines represent overall intensity levels from 0 to 5000 Hz. The lines labeled with squares represent the results of linear regression analyses on selected frequency regions.

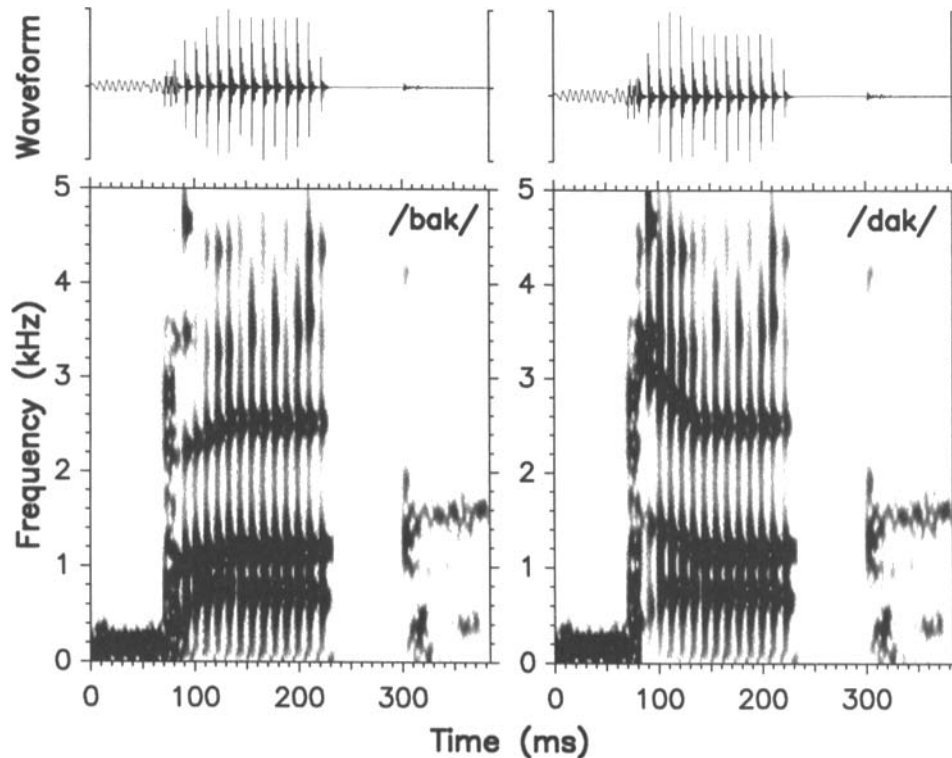


FIGURE 3. Waveform and spectrogram of the resynthetic endpoint tokens /bak/ (left panel) and /dak/ (right panel).

determines the value of the phoneme boundary and slope by iteratively computing the cumulative normal distribution that comes closest to the data, using a maximum likelihood criterion. The resulting distribution has a mean (i.e., the interpolated 50% crossover point or phoneme boundary) and a standard deviation (i.e., a measure of the variability of scores around the mean). The slope is the reciprocal of the standard deviation and indicates the range of uncertainty in distinguishing one phoneme category from another. A high slope value indicates a small uncertainty range and suggests a highly consistent ability to categorize a speech contrast, whereas a low slope value indicates a large range of uncertainty and suggests difficulties in identifying the speech stimuli. In Table 2, the mean phoneme boundary and slope scores for the apraxic and the control groups are presented.

A two-factor analysis of variance (unbalanced design) was used to test for differences in the scores. The two factors were Group (apraxic versus control) and Stimulus Type (resynthesis versus synthesis), with the levels of the latter factor treated as repeated measures.

A significant difference in the slopes for the resynthetic and synthetic continua [$F(1, 31) = 8.38, p < .01$] with a higher score on the resynthetic variant was found. The range of uncertainty was increased by the use of synthetic stimuli. The mean slope for the group with apraxia did not differ significantly from the mean slope for the control group [$F(1, 31) = 0.02, p = .89$], showing that the children with apraxia performed as consistently as the control children. Also not found was a significant interaction between Group and

Stimulus Type [$F(1, 31) = 0.33, p = .57$], indicating that the synthetic stimuli did not improve our ability to distinguish children with apraxia from control children.

The phoneme boundary for the group with apraxia compared to the control group was significantly shifted to the left [$F(1, 31) = 6.68, p < .05$]. Children with developmental apraxia demonstrated a bias for the perception of /d/. The phoneme boundary for the resynthetic continuum when compared to the synthetic continuum was also shifted to the left [$F(1, 31) = 10.72, p < .01$]. Subjects identified resynthetic more than synthetic stimuli as /d/. The latter effect was independent of Group, as indicated by the lack of an interaction between Group and Stimulus Type [$F(1, 31) = 0.00, p = .9725$].

Discrimination

The discrimination functions, which plot the discriminability of each 2-step pair in terms of a nonparametric estimate of d' (-ln eta, Wood, 1976) as a function of stimulus pair, are shown in Figure 6 for the resynthetic and synthetic pairs, respectively. Discriminability equals zero when performance is at chance. It increases with increasing discrimination accuracy without influences of bias to respond "same" or "different." Discriminability is maximal at the -ln eta value of 4.6; this value is obtained when the probabilities for the correct "difference" and correct "same" responses are both .99, the value assigned (for computational purposes) when the actual probabilities were 1.00.

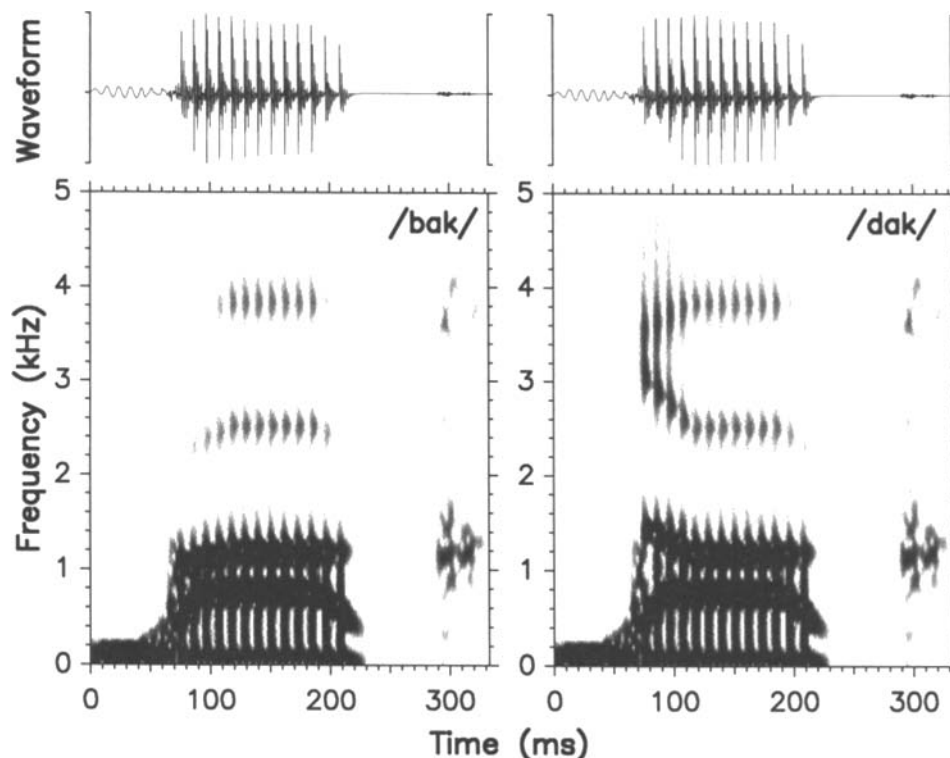


FIGURE 4. Waveform and spectrogram of the synthetic endpoint tokens /bak/ (left panel) and /dak/ (right panel).

A multivariate analysis of variance on the discriminability data (MANOVA, Group \times Stimulus Type, statistic: Wilks' lambda) revealed a significant effect of Group [$F(5, 27) = 2.8623, p < .05$]. The mean discrimination pattern for the children with apraxia was different from that of the control children. In addition, there was a Stimulus Type difference [$F(5, 27) = 2.7625, p < .05$], indicating differential treatment of the resynthetic versus synthetic stimulus pairs. The extent of the stimulus type difference was, moreover, dependent on Group [interaction Group and Stimulus Type: $F(5, 27) = 3.7261, p < .05$].

In order to specify the general effects in the MANOVA, a 3-way ANOVA (Group \times Stimulus Type \times Stimulus Pair) with repeated measures on Stimulus Type and Stimulus Pair was performed. Identification results demonstrated that the mean phoneme boundary for the children with apraxia was shifted to the left. Shifts in the mean discrimination peak, as a mere result of shifts in phonemic identification, were confirmed by the significant Group \times Stimulus Pair interaction [$F(4, 124) = 4.69, p < .01$]. In addition, and more interesting, the significant differences between the children with apraxia and control children (results MANOVA) could be

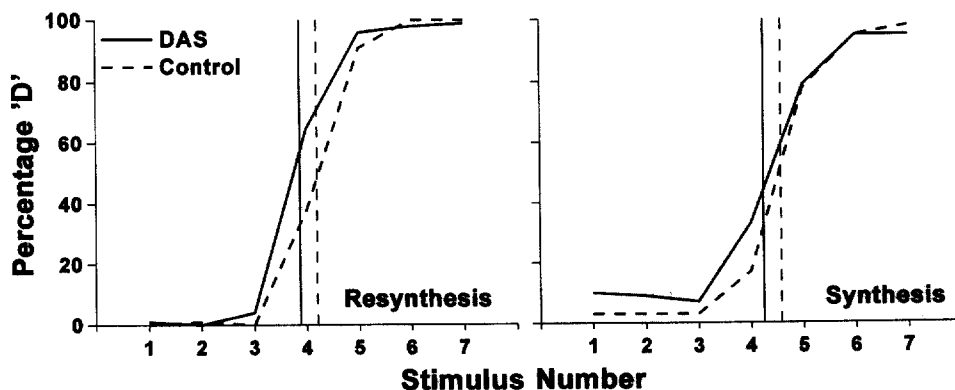


FIGURE 5. Mean percentages 'D' responses as a function of stimulus number for the children with apraxia (DAS) and control children in the resynthetic condition (left panel) and the synthetic condition (right panel). Vertical lines indicate the location of the mean phoneme boundary.

TABLE 2. Mean identification results for the children with apraxia and control children on the resynthetic and synthetic stimuli.

	Phoneme boundary	Slope
Apraxic		
Resynthesis	3.871	2.843
Synthesis	4.240	1.997
Control		
Resynthesis	4.196	2.667
Synthesis	4.573	2.102

accounted for by their overall discrimination levels: the children with apraxia performed more poorly [$F(1, 31) = 4.44, p < .05$]. Using synthetic speech did not produce a different pattern of responses for the two groups (no significant Group \times Stimulus Type interaction, $F(1, 31) = 1.50, p = .22$).

The children with apraxia demonstrated the same amount of uncertainty on perceptual categorization in the region of the phoneme boundary as the control children. This suggests that phonetic processing was not affected and that the poorer overall discrimination performance of the children with apraxia is the result of deficient auditory processing.

Discussion

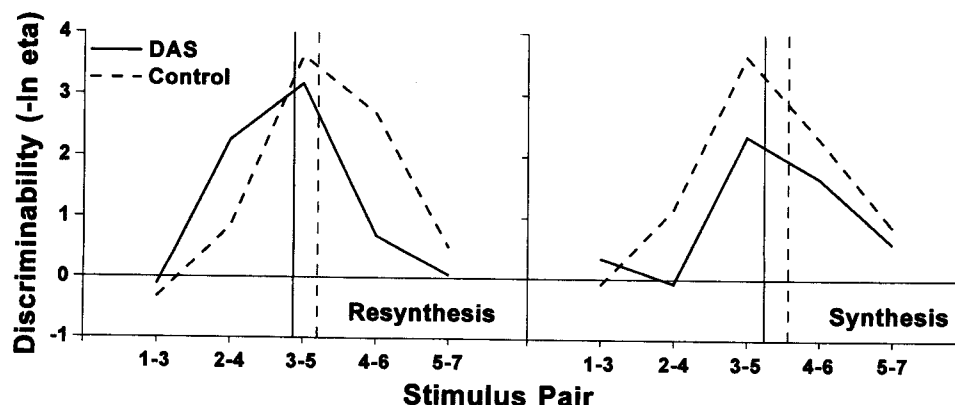
The results of Experiment 1 showed children with apraxia to have equally steep identification slopes when compared to control children, indicating equally consistent phonetic processing across groups. The mean phoneme boundary for the group with apraxia is shifted to the left when compared to the control group, suggesting that for this group the stage where acoustical properties are combined to form a phonemic judgment may be biased towards the alveolar variant. However, the extent to which the phoneme boundary is shifted never exceeds one stimulus on the stimulus dimension, explaining the lack of dramatic identification problems.

Perceptual differences between children with apraxia and control children were indicated by the significant differences

in their discrimination scores. The children with apraxia demonstrated poorer discrimination than the control children, which indicates poorer auditory processing. Children with apraxia appear to have less access to information in auditory memory than children without apraxia.

Intensity of the third formant (resynthesis versus synthesis) was found to influence both identification and discrimination. The identification and discrimination performance of both groups was poorer on synthetic than on resynthetic stimuli. However, the spectrally degraded synthetic stimuli hardly affected the performance of the children with apraxia more than the performance of the control children, which suggests insignificant clinical value. This finding complements those of Huntress et al. (1990) for aphasic subjects. An instant effect of synthetic speech was found along with improved speech comprehension by the aphasics with practice. The aphasic performance was not compared to that of a control group, however, so little could be said about the clinical value of the synthetic speech used in their study.

Reducing the intensity of the third formant in the synthetic speech tokens can be viewed as a reduction of the redundancy of the speech signal. The manipulation of intensity does not affect the integrity of the transiential character of the formants, and as such does not degrade aspects of the signal with high cue value for establishing the phonemic contrast. It can be concluded that this reduction of the redundancy of the speech stimulus in our study does not increase its clinical value or the sensitivity of perception tasks. Lowering the intensity of the third formant with the use of degraded synthetic speech may simply not tap into central auditory problems because of the potentially marginal relation to the speech signal and the specific processes of speech perception. Inducing overall spectral poorness (in the present study by lowering the intensity of the third formant) thus appears to have little value for the assessment of central auditory problems. In contrast, a reduction of the stimulus redundancy by using speech continua appears to have clinical value. This method of reduction pertains to dimensions that directly concern features essential for speech recognition (i.e., formant transitions in the perception of place-of-articulation contrasts) and therefore appears to have a direct effect on the process of speech perception.

**FIGURE 6.** Mean discrimination scores as a function of stimulus pair for the children with apraxia and control children in the resynthetic condition (left panel) and the synthetic condition (right panel). Vertical lines indicate the location of the mean phoneme boundary.

EXPERIMENT 2

In Experiment 2, the hypothesis regarding a relation between the speech-processing deficit in the children with apraxia and specific problems with regard to the production of place-of-articulation information was tested.

Method

Subjects

The subjects were the same 17 children with apraxia as in Experiment 1.

Materials

Two articulation tests of 30 words and 36 pseudowords (see Appendix A) were used in this experiment. Apart from semantics, the tests using real words had a higher phonemic complexity and a higher complexity of syllabic structure. The tests were previously used in a study of verbal apraxic articulatory behavior by Thoonen et al. (1994).

Procedure

Each subject was asked to imitate 30 single words and 36 single pseudowords. The repetitions were recorded by tape recorder and phonetically transcribed (IPA) by three certified speech-language pathologists.

Analysis

To assess the reliability of the phonetic transcriptions, an interjudge agreement analysis was carried out across the total set of 66 words (30 words and 36 pseudowords) for three randomly selected children with developmental apraxia of speech and three of the children with normal speech. The latter children were matched for age and gender to the three children with apraxia. The prime transcriber and two additional transcribers, each fully acquainted with the procedures of IPA transcription, transcribed a total of 1,542 consonants in 428 utterances (mean of 3.9 consonants per utterance). The Pearson correlation coefficients across the judges for the total number of one-feature errors (place, manner, and voicing) and multiple-feature errors per subject were between .92 and .99 ($p < .05$).

The transcriptions of the consonants in particular were then analyzed and transferred into confusion matrices. In matrices for each target feature (place, manner, and voice), the percentage of substituted consonants that were correct with respect to that particular feature (but in error with respect to one or both of the other features) was calculated. The "percentage feature retention" was tabulated for each target feature and—within each feature—for each value of the feature (e.g., target feature: place; values: bilabial, labiodental, alveolar, palatal, velar, glottal).

Five scores for place-of-articulation were then derived to

represent each child's performance. (a) The proportion "one feature" place-substitutions (e.g., target: /p/, realization: /t/) was calculated across the total number of initial syllabic consonants for words and for (b) pseudowords. (c) The proportion "multiple feature" place-substitutions including place-manner (e.g., target: /p/, realization: /s/), place-voice (e.g., target: /p/, realization: /d/), and place-manner-voice (e.g., target: /p/, realization: /l/) was calculated across the total number of initial syllabic consonants for words and for (d) pseudowords. Finally, (e) the mean proportion place-substitutions across all words and pseudowords (i.e., the number of "one-feature" place-substitutions and "multiple feature" place-substitutions divided by the total number of consonants in the words and pseudowords) was calculated. In order to determine the specificity between place perception and place production, we also calculated these five production scores for manner-of-articulation and voicing.

Experiment 1 had shown the children with apraxia to discriminate the place contrast less well than the control children. The production scores for place-of-articulation were therefore compared to the mean discriminability of the resynthetic stimuli and the synthetic stimuli in order to examine the specific relation between perception and production.

Results

In Table 3, the mean scores for the five production measures and two perception measures for the children with apraxia are presented.

In Figure 7, the mean discriminability for place-of-articulation is plotted against the mean proportion place substitutions per subject for both the resynthetic and synthetic conditions. Cross-correlation of the production variables with the perception variables produced the Pearson product-moment correlation coefficients as listed in Table 4. If perception is related to production (i.e., high mean discriminability level and low proportion place-substitutions), then the correlations should be significantly negative.

The mean discriminability scores for the resynthetic and synthetic conditions were negatively related to the mean proportion place substitutions ($r = -.56, p < .05$ and $r = -.61, p < .01$, respectively). The mean discriminability of the resynthetic condition was negatively related to the proportion "one feature" place-substitutions for both words and pseudowords ($r = -.49, p < .05$ and $r = -.54, p < .05$, respectively). The mean discriminability of the synthetic condition was negatively related to the proportions of "multiple feature" place-substitutions for both words and pseudowords ($r = -.50, p < .05$ and $r = -.64, p < .01$, respectively). In addition, the synthetic condition was negatively related to the proportion "one feature" place-substitutions in words ($r = -.49, p < .05$). There were no significant correlations between the perception scores and production substitutions for the manner-of-articulation or voicing.

The correlation coefficients revealed a specific relation between place perception and place production. In order to

TABLE 3. Means and standard deviations for the production and perception scores for the children with apraxia.

	Mean	Standard deviation
Production (proportion place substitutions)		
"One Feature"		
words	.058	.064
pseudowords	.082	.029
"Multiple Feature"		
words	.045	.041
pseudowords	.054	.027
Mean Proportion	.061	.029
Perception (mean discriminability)		
Resynthesis	1.210	0.697
Synthesis	0.981	0.976

examine the utility of this relation for the diagnosis of children with apraxia, the relation between the perception and production scores for only the poorest performers was examined. Those children with moderate to severe developmental apraxia of speech were selected out of the group of 17 children with apraxia. This selection was based on the perception as well as the production data and produced a number of different subject subsets. The overlap between these subsets was then examined to determine the number of children who performed particularly poorly on both perception and production.

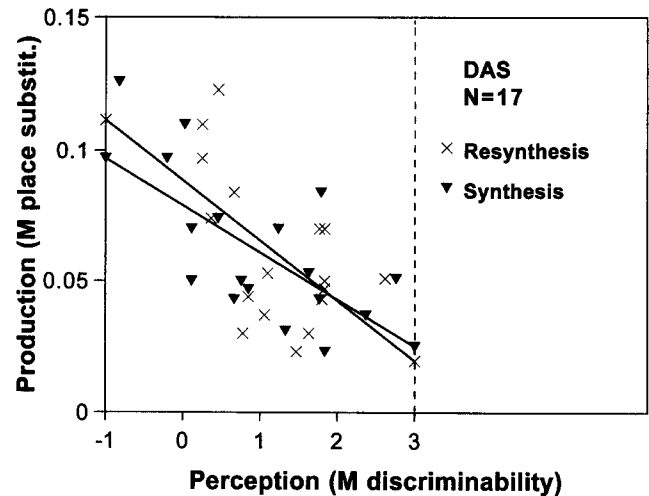
An objective division based on the number of place-substitutions was made. Randomization testing (Edgington, 1987) was used to nonparametrically single out the poorest performing cases with apraxia. Exhaustive randomization testing was performed on the four proportion measures for place-substitutions (the mean total proportion place-substitutions was excluded), and all of the possible solutions for the selection of a subgroup of six children were tested. This randomization test identified the six poorest performing subjects with apraxia (the test statistic used was equivalent to F for MANOVA, see details of the selection process in Appendix B).

In selecting the poorest performing cases with apraxia on perceptual grounds, the mean discriminability of (a) the resynthetic stimuli and (b) the synthetic stimuli were taken into consideration. For each of the two measures, the six children with the poorest scores were selected for comparison.

Comparison of the different poorest performing subsets produced overlap frequencies of 5 for both the resynthetic and the synthetic condition. From the six poorest performing subjects in perception, five were included in the six poorest performing subjects in production.

The significance levels for the overlap were calculated using the following formula:

$$p = \frac{\binom{S}{O} * \binom{N-S}{S-O}}{\binom{N}{S}} \quad (1)$$

**FIGURE 7.** Scatterplot of the individual perception and production scores for place-of-articulation. Regression lines are included for both the resynthetic and synthetic conditions.

where

N = Total number of children with apraxia (= 17)

S = Number of children in selection (= 6)

O = Number of matches (= 5)

and binomial coefficients, e.g.:

$$\binom{S}{O} = \frac{S!}{O!(S-O)!} \quad (2)$$

where

$S! = 1*2*3*...*(S-1)*S$

etcetera.

If $O \geq 5$ (i.e., at least five matches), the cumulative p is .005. This indicates an O -value significantly higher than that expected with random selection.

Comparison of the subsets based on the perceptual discriminability scores and the subsets based on the randomization testing procedure for production produced a significant number of matches for both the resynthesis and synthesis conditions. In addition, of the five matches in the resynthetic conditions, four were included in the five matches in the synthetic condition, which indicates the reliability of the perceptual tasks.

The results of the poorest performing subset comparisons emphasize the relation between problems in perception and problems in production. The overlap between the poorest performing subjects in perception and the poorest performing subjects in production was found to be large, which strongly indicates the clinical value of the discrimination tasks.

Discussion

In Experiment 2, the specific relation between the perception and production of place-of-articulation was studied. A

TABLE 4. Pearson product-moment correlation coefficients for the five production measures with the two perception measures for place, manner, and voicing.

	Proportion substitutions				
	"One Feature"		"Multiple Feature"		Mean proportion
	words	pseudo	words	pseudo	
Place subst.					
Mean Discriminability Place					
Resynthesis	-.49*	-.54*	-.36	-.27	-.56*
Synthesis	-.49*	-.22	-.50*	-.64**	-.61**
Manner subst.					
Mean Discriminability Place					
Resynthesis	.31	-.32	-.07	-.12	-.14
Synthesis	-.22	-.46	-.03	-.21	-.39
Voicing subst.					
Mean Discriminability Place					
Resynthesis	-.00	.33	-.01	-.09	.21
Synthesis	-.19	.29	.27	-.22	.15

* $p < .05$ ** $p < .01$

feature-to-feature comparison was undertaken. The degree of disturbance of place discrimination was found to be closely related to the number of place substitutions in speech production, with no relation between place discrimination and manner or voicing substitutions in production. This supports the feature-specific interdependence of perception and production. The clinical significance of this interdependence was also demonstrated by comparing the poorest performing subjects in production to the poorest performing subjects in perception, which resulted in a significant number of matches. In addition, the results of our experiment cross-validate the methods used to assess perception and production.

GENERAL DISCUSSION

At the theoretical level, our results pose some interesting problems. According to a dual-coding structure of auditory and phonetic processing, it is difficult for auditory discrimination to be disturbed when phonetic identification is normal. That is, dual-coding models of speech perception as described in Sawusch and Nusbaum (1983) and Sawusch and Mullenix (1985) do not provide a clear rationale for this phenomenon. The question, then, is whether or not consistent phonetic identification can develop with problems of auditory processing. One can assume task or stimulus factors to be a possible explanation for our results, but these were ruled out in the present study. One cannot question the materials (e.g., the use of ecologically invalid stimuli) or the procedure (e.g., the use of "same/different" tasks). Rather, a substantial, feature specific, perceptual component in developmental apraxia of speech appears to exist. This means that the validity of classical models of speech perception should be questioned. A better mapping process from the acoustic to phonetic stages seems needed, one that may depend on language exposure. Because discrimination performance was affected whereas identification consistency was not, the results of our study support a structure for

speech processing with an auditory stage and a phonetic stage partly allowing for stage-independent output and the integrity of phonetic processing not being totally dependent on the outcome of auditory processing. This view is supported in a theoretical note from Ades (1977) on the perception of speech and nonspeech. He suggests the possibility of phonetic processing not receiving input from acoustical traces and instead forming an entirely independent route.

Another explanation for the incongruity of our findings with a hierarchical model of speech perception could lie in subject strategy factors. If the dysfunctional auditory profile of children with developmental apraxia of speech is assumed to be a reflection of the application of an abnormal discriminative strategy, then it is not unlikely that the auditory dysfunction may be the result of selective attention. Nittrouer and Studdert-Kennedy (1987) and Nittrouer (1992) demonstrated that children show different cue weighting strategies than adults. In studies on the perception of syllable-initial fricatives, it was shown that preschool and young school-age children weighted intrasyllabic formant transitions more than adults did in making phonetic decisions. Dysfunctional auditory processing, then, may be determined by the process of attributing different (yet valid and adequate) weights to acoustic cues in the process of categorization (phonetic processing) as compared to the cues on which auditory discrimination was based. This would allow discrimination performance to be affected although phonetic identification was not (one of the results of this study). Such a viewpoint is highly compatible with the fuzzy logical model of speech perception from Massaro (1987, 1992). This viewpoint also seems compatible with the concepts of the WRAPSA model (Word Recognition and Phonetic Structure Acquisition) of Jusczyk, 1993). According to Jusczyk, preliminary auditory analysis reflects the inherent organization of the human auditory system. The auditory analyzers belong to the innate endowment of the infant. After the development of a weighting scheme, language-specific phonetic decisions can be made. Selective

attention within a set of intraphonetic cues, then, seems likely to play a role in phonetic processing. It is possible that for children with DAS, cues playing a minor role are weighted less in phonetic processing. Thus, problems with auditory processing of these minor cues may not have as much importance for phonetic processing as might problems in discrimination of more important speech cues, such as the second formant transition for perception of place-of-articulation.

An interesting issue in trying to understand developmental apraxia of speech is whether the underlying dysfunction should be sought at the level of output motor processes or at the level of phonological representations. Marion et al. (1993) found children with apraxia to have rhyming abilities inferior to those of children without apraxia. They suggest a conceptualization of DAS as a fundamental disorder of the segmental phonological level of language that influences all relevant language components. Just as interesting as the etiology of DAS, and perhaps related, is the potential value of perceptual tasks in the differentiation of psycholinguistic pathologies. Using tasks similar to those in the present study, Sussman (1993) found children with language impairments to show poorer phonetic categorization than children without language impairments. Her results support hypotheses from Gathercole and Baddeley (1990) suggesting that the phonological component of working memory may be disordered in children with language impairment. Should developmental apraxia of speech be a dysfunction at an output motor level, then there may be a specific symptomatology associated with the dysfunction and different from that for children with language impairment. We, indeed, found children with apraxia to have problems in discrimination (i.e., auditory processing) and not in phonetic categorization. A straightforward conclusion is that perceptual tasks have differential clinical value. Moreover, our results appear to be incompatible with an explanation at the level of phonological representation and appear to favor the conception of DAS as a senso-motoric problem (see MiCoCh & Square, 1984).

A different and difficult problem arises in the determination of the causal connection between perception and production. Studies of normal perception and production skills have shown many instances of perception preceding production. In reviewing the potential causal relation between perception and production, Strange and Broen (1980) point out that human infants appear to be sensitive to acoustic-phonetic dimensions that allow them to perceive phoneme contrasts far in advance of any ability to produce such contrasts. Kuhl (1991) has shown that phonetic category prototypes exist at an age of 6 months, serving as language-specific "perceptual magnets" for other stimuli. Her results have led to the development of the Native Language Magnet (NLM) theory, which describes how innate factors along with experience with a specific language form the development of speech perception (Kuhl, 1993). The earlier-mentioned WRAPSA model (Jusczyk, 1993) assumes the emphasis on the critical dimensions needed for distinguishing among words in the native language to be developed during the second half of the first year of life. Whereas the WRAPSA (Jusczyk, 1993)

and the NLM (Kuhl, 1993) differ in emphasis on lexical/phonological contrast and phonetic categories, respectively, both theories suggest that during the first year of life, before infants acquire word meaning and contrastive phonology and before the critical time for the development of production skills, essential phonetic perception strategies have been developed. In keeping with this view, Rvachew and Jamieson (1989) report that certain speech perception deficits may in fact contribute to articulation disorders.

In addition to the idea that perception precedes production, however, there are also signs that production can affect perception (Crul, 1990; Monnin & Huntington, 1974). Hoffman et al. (1985) hypothesize that children with certain articulation problems may not experience enough communicative pressure to develop the critical phonological contrasts in sufficient phonetic detail. Their productive neutralization may then lead them to ignore the contrast perceptually as well.

The present data do not address the causal relation between perception and production directly. It is likely that the relevant phonological processes in perception and production interact. The presumed neurological defect that initially disrupted the oral-motor functioning of the child may also have disrupted the auditory functioning of the child at the same time. In any case, the present study clearly shows auditory perception of children with developmental apraxia to be affected, and this fact must be considered not only in the diagnosis of developmental apraxia of speech but also in the development of a valid course of treatment.

Acknowledgments

The Netherlands Organization for Scientific Research (NWO) is gratefully acknowledged for funding this project. This research was conducted while Paul Groenen was supported by a PSYCHON-grant from this organization (560-256-047) awarded to Dr. Ben Maassen and Dr. Thom Crul. We would also like to thank Henk Loman for help in the construction of the synthetic stimuli and Henriët Ruuls and Jan Wit for their cooperation in the testing of the children. Finally, we want to thank the children and teachers of the special schools Martinus van Beekschool (Nijmegen), Dr. Bossschool (Arnhem), De Horst (Eindhoven), and Mgr. Terwindtschool (Groesbeek) and the regular school De Piekenstulp (Wijchen) for their cooperation.

References

- Ades, E. A. (1977). Vowels, consonants, speech, and nonspeech. Theoretical note. *Psychological Review*, 84, 524-530.
- Baddeley, A. (1992). Working memory. *Science*, 255, 556-559.
- Boves, L., Kerkhoff, J., & Loman, H. (1987). A new synthesis model for an allophone based text-to-speech system. In J. Laver & M. A. Jack (Eds.), *Proceedings of the European Conference on Speech Technology*, Edinburgh (pp. 385-388). Edinburgh.
- Bridgeman, E., & Snowling, M. (1988). The perception of phoneme sequence: A comparison of dyspraxic and normal children. *British Journal of Disorders of Communications*, 23, 245-252.
- Coombs, C. H., Dawes, R. M., & Tversky, A. (1970). *Mathematical psychology*. Englewood Cliffs, NJ: Prentice-Hall, Inc.
- Crory, M. A. (1984). Phonological characteristics of developmental verbal dyspraxia. In D. M. Aram (Ed.), *Seminars in speech and*

- language: Assessment and treatment of developmental apraxia (pp. 71-83). New York: Thieme-Stratton Inc.
- Crary, M. A., Landess, S., & Towne, R. (1984). Phonological error patterns in developmental verbal dyspraxia. *Journal of Clinical Neuropsychology*, 6, 157-170.
- Crul, Th. A. M. (1990). De spraakperceptie-spraakproductielink omgekeerd [Reversing the speech perception-speech production link]. *Van Horen Zeggen*, 2, 62-69.
- Cutting, J. E., & Pisoni, D. B. (1978). An information-processing approach to speech perception. In J. F. Kavanagh & W. Strange (Eds.), *Speech and language in the laboratory, school and clinic* (pp. 38-87). Cambridge, MA: The MIT Press.
- Edgington, E. S. (1987). *Randomization Tests*. New York: Marcel Dekker.
- Edwards, M. (1984). *Disorders of articulation* (pp. 77-97). New York: Springer Verlag.
- Ekelman, B. L., & Aram, D. M. (1983). Syntactic findings in developmental verbal apraxia. *Journal of Communication Disorders*, 16, 237-250.
- Finney, D. J. (1971). *Probit analysis*. Cambridge: University Press.
- Gathercole, S., & Baddeley, A. (1990). Phonological memory deficits in language disordered children: Is there a causal connection? *Journal of Memory and Language*, 29, 336-360.
- Guyette, T. W., & Diedrich, W. M. (1981). A critical review of developmental apraxia of speech. In N. J. Lass (Ed.), *Speech and language: Advances in basic research and practice* (pp. 1-47). New York: Academic Press.
- Hall, P. K. (1992). *Developmental apraxia of speech: Cases and questions*. Paper presented at the Annual Convention of the American Speech-Language-Hearing Association (ASHA). San Antonio, TX, November 20-23.
- Hall, P. K., Jordan, L. S., & Robin, D. A. (1993). *Developmental apraxia of speech*. Austin, TX: Pro-Ed.
- Hammarberg, B., Fritzell, B., Gauffin, J., Sundberg, J., & Wedin, L. (1980). Perceptual and acoustic correlates of abnormal voice qualities. *Acta Otolaryngologica*, 90, 441-451.
- Hoffman, P. R., Daniloff, R. G., Bengoa, D., & Schuckers, G. H. (1985). Misarticulating and normally articulating children's identification and discrimination of synthetic /r/ and /w/. *Journal of Speech and Hearing Disorders*, 50, 46-53.
- Holt-Dalgaard, J., Murry, T., & Kopp, H. G. (1983). Voice onset time production and perception in apraxic subjects. *Brain and Language*, 20, 329-339.
- Huntress, L. M., Lee, L., Creaghead, N. A., Wheeler, D. D., & Braverman, K. M. (1990). Aphasic subjects' comprehension of synthetic and natural speech. *Journal of Speech and Hearing Disorders*, 55, 21-27.
- ILS-PC, V6.1. (1989). *Interactive Laboratory System*. Goleta: Signal Technology, Inc.
- International Organization for Standardization. (1985). *ISO Recommendation R 389: Standard reference zero for the calibration of pure-tone audiometers* (2nd ed.). Switzerland: ISO.
- Jusczyk, P. W. (1993). From general to language-specific capacities: The WRAPSA model of how speech perception develops. *Journal of Phonetics*, 21, 3-28.
- Kuhl, P. K. (1991). Human adults and human infants show a "perceptual magnet effect" for the prototypes of speech categories; monkeys do not. *Perception and Psychophysics*, 50, 93-107.
- Kuhl, P. K. (1993). Early linguistic experience and phonetic perception: Implications for theories of developmental speech perception. *Journal of Phonetics*, 21, 125-139.
- Liberman, A. M., Harris, K., Hoffman, H. S., & Griffith, B. C. (1957). The discrimination of speech sounds within and across phoneme boundaries. *Journal of Experimental Psychology*, 54, 358-368.
- Locke, J. L. (1980a). The inference of speech perception in the phonologically disordered child. Part I: A rationale, some criteria, the conventional tests. *Journal of Speech and Hearing Disorders*, 45, 431-444.
- Locke, J. L. (1980b). The inference of speech perception in the phonologically disordered child. Part II: Some clinically novel procedures, their use, some findings. *Journal of Speech and Hearing Disorders*, 45, 445-468.
- Loman, H., Kerkhoff, J., & Boves, L. (1989). A working environment for speech synthesis by rule. *Proceedings of the Institute of Phonetics, University of Nijmegen*, 13, 51-63.
- Marion, M. J., Sussman, H. M., & Marquardt, T. P. (1993). The perception and production of rhyme in normal and developmentally apraxic children. *Journal of Communication Disorders*, 26, 129-160.
- Markel, J. D., & Gray, A. H. (1976). *Linear prediction of speech*. Berlin-Heidelberg: Springer-Verlag.
- Marquardt, T. P., Dunn, C., & Davis, B. (1985). Apraxia of speech in children. In J. Darby (Ed.), *Speech and language evaluation in neurology* (pp. 113-129). New York: Grune & Stratton, Inc.
- Massaro, D. W. (1987). Integrating multiple sources of information in listening and reading. In A. Allport, D. Mackay, W. Prinz, & E. Scheerer (Eds.), *Language perception and production: Shared mechanisms in listening, speaking, reading, and writing* (pp. 111-129). London: Academic Press.
- Massaro, D. W. (1992). Broadening the domain of the fuzzy logical model of speech perception. In H. L. Pick, P. van den Broek, & D. C. Kroll (Eds.), *Cognition: Conceptual and methodological issues* (pp. 51-84). Washington DC: American Psychological Association.
- McCoch, A. G., & Square, P. A. (1984). Apraxia of speech: Articulatory and perceptual factors. In N. Y. Lass (Ed.), *Speech and language: Advances in basic research and practice* (Vol. 10, pp. 1-57). New York: Academic Press Inc.
- Monnin, L. M., & Huntington, D. A. (1974). Relationship of articulatory defects to speech-sound identification. *Journal of Speech and Hearing Research*, 17, 352-366.
- Nittrouer, S., & Studdert-Kennedy, M. (1987). The role of coarticulatory effects in the perception of fricatives by children and adults. *Journal of Speech and Hearing Research*, 30, 319-329.
- Nittrouer, S. (1992). Age-related differences in perceptual effects of formant transitions within syllables and across syllable boundaries. *Journal of Phonetics*, 20, 351-382.
- Ohde, R. N., & Sharf, D. J. (1988). Perceptual categorization and consistency of synthesized /r-w/ continua by adults, normal children and /r/-misarticulating children. *Journal of Speech and Hearing Research*, 31, 556-568.
- Pisoni, D. B., & Sawusch, J. R. (1975). Some stages of processing in speech perception. In A. Cohen & S. G. Nooteboom (Eds.), *Structure and process in speech perception* (pp. 16-34). Berlin-Heidelberg-New York: Springer-Verlag.
- Raaymakers, E. M. J. A., & Crul, Th. A. M. (1988). Perception and production of the final /s-ts/ contrast in Dutch by misarticulating children. *Journal of Speech and Hearing Disorders*, 53, 262-270.
- Repp, B. H. (1984). Categorical perception: Issues, methods, findings. In N. J. Lass (Ed.), *Speech and language: Advances in basic research and practice* (pp. 243-335). New York: Academic Press.
- Rvachew, S., & Jamieson, D. G. (1989). Perception of voiceless fricatives by children with a functional articulation disorder. *Journal of Speech and Hearing Disorders*, 54, 193-208.
- Sawusch, J. R., & Mullennix, J. W. (1985). When selective adaptation and contrast effects are distinct: A reply to Diehl, Kluender, and Parker. *Journal of Experimental Psychology: Human Perception and Performance*, 11, 242-250.
- Sawusch, J. R., & Nusbaum, H. C. (1983). Auditory and phonetic processing in place perception for stops. *Perception and Psychophysics*, 34, 560-568.
- Snowling, M., & Stackhouse, J. (1983). Spelling performance of children with developmental verbal dyspraxia. *Developmental Medical Child Neurology*, 25, 430-437.
- Stackhouse, J. (1992). Developmental verbal dyspraxia I: A review and critique. *European Journal of Disorders of Communication*, 27, 19-34.
- Strange, W., & Broen, P. A. (1980). Perception and production of approximant consonants by three-year-olds: A first study. In B. H. Yeni-Komshian, J. F. Kavanagh, & C. A. Ferguson (Eds.), *Child phonology: Perception* (Vol. 2, pp. 117-154). New York: Academic Press.
- Studdert-Kennedy, M., Liberman, A., Harris, K., & Cooper, F. (1970). Motor theory of speech perception: A reply to Lane's critical review. *Psychological Review*, 77, 234-249.

- Sussman, J. E.** (1993). Perception of formant transition cues to place of articulation in children with language impairments. *Journal of Speech and Hearing Research*, 36, 1286–1299.
- Thoonen, G., Maassen, B., Gabreëls, F., & Schreuder, R.** (1994). Feature analysis of singleton consonant errors in developmental verbal dyspraxia (DVD). *Journal of Speech and Hearing Research*, 37, 1424–1440.
- Thoonen, G., Maassen, B., Wit, J., Gabreëls, F., & Schreuder, R.** (in press). The integrated use of maximum performance tasks in differential diagnostic evaluations among children with motor speech disorders. *Clinical Linguistics and Phonetics*.
- Wood, C. C.** (1976). Discriminability, response bias, and phoneme categories in discrimination of voice onset time. *Journal of the Acoustical Society of America*, 60, 1381–1389.

Received February 22, 1995

Accepted January 10, 1996

Contact author: Paul Groenen, University Hospital Nijmegen, Department of Otorhinolaryngology, P.O. Box 9101, 6500 HB Nijmegen, The Netherlands. E-mail: kno_pg@aznvx1.azn.nl

Appendix A

Contents of lists used to test articulatory skills

List 1 Words (30 items)

1 nijmegen	11 fopspeen	21 koekoeksklok
2 diploma	12 fietstas	22 voetbalwedstrijd
3 pyjama	13 kokskind	23 kleurpotlood
4 limonade	14 kapstok	24 bladzijde
5 meenemen	15 nepski	25 spiksplinternieuw
6 liniaal	16 tikspel	26 tandpasta
7 portemonnee	17 kletskous	27 gereedschapkast
8 paraplu	18 prikstok	28 hartstikkestoer
9 lokomotief	19 hutspot	29 bliksemflits
10 paddestoel	20 viltstift	30 brandweerspuit

List 2 Pseudowords (36 items)

1 pabata	13 tadaka	25 nama
2 faxaga	14 xafaka	26 bapa
3 ladata	15 mabapa	27 lana
4 katada	16 pataba	28 daba
5 nalama	17 lanama	29 data
6 bapama	18 tadana	30 nala
7 bakaba	19 pata	31 paka
8 saxata	20 bapa	32 paba
9 tanada	21 lada	33 sata
10 bapaka	22 xafa	34 safa
11 xatasa	23 tada	35 nada
12 dalata	24 kaxa	36 tada

Appendix B

Randomization test

Randomization testing (see Edgington, 1987) is intended to determine nonparametrically the probability value associated with a given comparison. The statistic used to test the difference between the two functions is:

$$\sum_{j=1}^s \sum_{i=1}^k T_{ij}^2 / N_i$$

where

T_{ij} = total score of group i and variable j

N_i = number of subjects in group i

k = number of groups being compared (in our case 2)

s = number of variables (in our case 4)

The resulting test statistic is equivalent to the F of MANOVA. The probability value associated with the computed test statistic is determined by randomization test as follows. The data from the subjects in the groups being compared are repeatedly divided into groups. The data for each subject consist of the vector of scores on each variable. The vectors are permuted as a unit. The test statistic is computed for each of 1,000 randomly selected permutations. The probability value is the proportion of the test statistics equal to or greater than the test statistic obtained from the actual division of subjects into groups.

Increasing F s reflect increasing differences between groups. In the present study, we wanted to select the seven poorest performing subjects based on four variables. In other words, we wanted a division into groups with the largest possible test statistic. This subgroup of seven could be computed by performing randomization tests on the four variables. In order to make an exhaustive search for the best division, the number of permutations was not fixed to 1,000. Our program tested all possible solutions for selecting a subgroup of 6 out of 17 children (12,376 permutations). The division associated with the highest test statistic consisted of the six poorest performing subjects.